**Technical Case Study**

*Coding and Problem-Solving Skills*

This interview will be divided into 2 sections. Section 1 should be completed prior to the interview. You should be prepared to share your screen to walk through your approach as defined in the “Your Task” portion of Section 1. Section 2 will be an open-ended discussion during the interview time slot.

# **Section 1: Coding Skills**

## Instructions

Please read through the prompt below. Using the attached data, please answer the following questions using either Python or SQL, whichever you prefer. The objective of this exercise is to assess both your problem solving and coding skills. To that end, please keep in mind code efficiency and interpretability, as well as how best to address each problem to answer the specific stakeholder’s needs. As part of the interview, you will be asked to walk through your codelines and associated output. We will ask you questions about various decisions you made (both technical and non-technical) along the way.

## Data Inputs

Attached in your email, you will find the data you can use for this case study. In this, you will find 4 data files and 1 data glossary. Below is a summary of each data file:

* **transaction\_item.csv:** This is the fact data. It contains details about which products were contained in each basket and relevant details about those baskets, such as which store they were transacted in, when the transaction occurred, which card was used to transact, and the value of the products.
* **prod\_dim.csv:** This is dimension data about products. It contains product descriptions as well as the product hierarchy for the retailer. The primary key between this table and the transaction table is prod\_id.
* **store\_dim.csv:** This is dimension data about stores. It contains information about which banners and formats the stores belong to. The primary key between this table and the transaction table is store\_id.
* **card\_dim.csv:** This is dimension data about customers. It contains information about which people own which cards. The primary key between this table and the transaction table is card\_id.

## Your Task: Customer Segmentation

**Context:** The merchants at our retailer want to understand customer loyalty. They plan to use this to determine if customers who buy certain groups of products are more loyal to the retailer than customers who buy other groups of products. This difference will be used to make strategic decisions around which categories to over- and under-invest in.

**Your Task:** Develop a customer-level segmentation which classifies each customer (prsn\_id) into 1 of ‘n’ segments. Each segment should represent varying levels of loyalty to the retailer. As part of your answer, be prepared to answer questions such as:

* How did you choose the number of segments (‘n’)?
* What metrics did you choose to determine loyalty? Why?
* In what ways might you utilize this segmentation to inform business strategy?

# **Section 2: Problem-Solving Skills**

In this section, we will present you with a real-world, retail-oriented business problem which can be solved through various data science techniques. We will ask you to talk through how you might think about the problem, what approach(es) you might use to get to a solution, and how you would present it to your stakeholder.